Fixing the controlled substance quota system: Why the law and the regulator must change

Marta E. Wosińska

Introduction

Resilience has become the watchword of American drug policy. Policymakers across the political spectrum now speak the language of buffer stock, redundancy, and surge capacity—a supply chain's ability to absorb a shock without leaving patients short. Resilience has become a stated federal priority, with executive orders directing agencies to prioritize domestic production, a Food and Drug Administration (FDA) Essential Medicines List meant to focus that effort, a 2025 order to stockpile a six-month supply of active ingredients for dozens of critical drugs, and a run of congressional hearings and proposed bills. Most of the push is aimed at geopolitical risk, especially exposure to China.

One group of drugs stands out against that agenda: controlled stimulants and opioids. Stimulants for ADHD, relied on by millions of patients, have been chronically hard to obtain since October 2022. Injectable opioids listed on the FDA Essential Medicines List—morphine, hydromorphone, and fentanyl—have been in shortage since 2017.

These shortages hit patients who cannot wait or go without. When injectable opioids run short, hospitals ration and substitute during surgery, childbirth, and end-of-life care. When stimulants run short, adults and children with ADHD are left unable to fill the prescriptions they depend on to work and learn.

What sets the controlled substance shortages apart is that a rigid government system turns ordinary disruptions into lasting shortages. The current push for resilience targets foreign vulnerability, but here the constraint is domestic and self-imposed—the Drug Enforcement Administration's (DEA) quota system, too rigid to adjust when demand rises or supply falters. That rigidity runs through every level of the system: the aggregate ceiling DEA sets, how the ceiling is subdivided, when manufacturers receive their share, and how much they may hold in inventory—all connected only loosely to what patients need.

Congress has held hearing after hearing on drug shortages and introduced a stack of bills, but none address the quota system behind the controlled substance shortages. Congress can fix this, not by telling DEA what to do but by shifting quota administration from DEA to FDA. DEA's discretion follows its enforcement mission, and no set of instructions can redirect it against the mission that shapes it; the fix is an agency whose mission is keeping medicine in supply.

This paper traces how DEA's administration of the quota system undermines resilience at every level, from the national ceiling to the allocations given to individual manufacturers. It then argues that these failures follow from DEA's enforcement mission rather than from anything supply requires, and that quota administration belongs to FDA—the agency that already approves these drugs, inspects the plants that make them, and answers for keeping them in supply. The move requires legislation, and the paper closes by laying out the statutory changes that would support it.

How the quota system undermines supply resilience

The quota system exists to hold controlled substance production to the country's medical need, a commitment rooted in international treaty and implemented through the Controlled Substances Act

(CSA). The idea is intuitive: cap how much can be made, and less can be diverted. But it is far harder to do well than the intuition suggests—and doing it badly leaves patients' medical needs unmet.

Congress built the system on two levels in the CSA: an aggregate ceiling on how much of each substance may be made nationally, and, beneath it, individual quotas parceling that ceiling out to manufacturers. The problems come from how DEA implements that structure. Its handling of the individual quotas hurts both injectable opioids and ADHD stimulants; for the stimulants, the aggregate ceiling compounds the harm.

Individual facility quota

DEA fragments each manufacturer's quota into multiple separate grants and often issues them late in the year, so usable quota can be stranded before production begins. The one buffer that could absorb a disruption, inventory, is capped in a way that can prevent a firm from making even the quota it holds. And when a disruption does hit, stranded quota cannot be moved to another manufacturer that is able to use it.

The fragmentation starts in the statute, which creates two quotas. The aggregate quota is a national ceiling on how much of a substance may be made; beneath it, a manufacturing quota authorizes each facility to produce up to a set quantity of an active ingredient. A third, the procurement quota, exists only in regulation and authorizes a facility to obtain and use a controlled ingredient as input. This means a plant making an active ingredient needs manufacturing quota to produce it and procurement quota to buy its controlled inputs, while a plant making the finished dose needs only procurement quota. DEA has since split these quotas into further subcategories.

Under DEA's current procedures, quota does not arrive once a year as the statute contemplates; it arrives in pieces. Most manufacturers receive only part of their allocation up front and must apply throughout the year for the rest, sometimes repeatedly, waiting months with little insight into how DEA decides what to grant or when. Across a few dozen controlled substances, these partial grants and repeated applications add up to roughly 4,000 quota decisions a year, each a point where supply can arrive late, fall short, or misalign with the stage that needs it.

Inventory is one buffer that can absorb a disruption, but the way DEA sets inventory rules can stop a manufacturer from using quota it has been granted. DEA limits how much of a controlled substance a facility may hold at any one time, counting every form it takes—raw material, work in progress, and finished product alike. The cap is set as a percentage of the facility's own recent sales, so a firm trying to produce more than it sold last year—a new generic entrant, say, or one stepping up to cover a competitor's shortage—may not be allowed to hold enough working material to make even the volume its quota already permits.

DEA administers the individual quotas in a way that strips the system of resilience. When production shuts down because of a contamination event or equipment breakdown—a recurrent, foreseeable hazard for these drugs—demand shifts to the others. DEA cannot prevent the disruption, and no regulator could make recovery instant. But the barriers that keep other manufacturers from stepping in—the inventory cap, the stranded quota, the quota that arrives too late—are DEA's, and they turn a recoverable disruption into a lasting shortage.

Aggregate market-wide quota

Every individual quota is carved from an aggregate ceiling, the national limit on how much of a substance may be made in a year; for the ADHD stimulants, that ceiling sits below what patients need. DEA is directed to size it to the country's estimated medical need, but it builds the estimate from past sales.

Past sales are a biased measure whenever supply has been constrained, and the bias feeds on itself. When quota is binding, sales cannot rise to meet growing demand because they reflect only what the quota allows. And when quota goes unused because the system could not get it into production, recorded sales can fall below even the prior year, which DEA then reads as lower need. Even when other signals point to rising demand—new diagnoses, shortage reports from patients and pharmacies—DEA is slow to correct the ceiling upward, treating some of it as misuse it should not accommodate. But whether a prescription is warranted is a judgment a production ceiling cannot make.

The aggregate ceiling is not the constraint for injectable opioids. DEA's own figures show injectable products using less than 5% of the relevant substance's quota, so the ceiling leaves ample room; the injectable failures sit below it, in the individual-quota problems already described.

Why DEA's enforcement mission trumps patient access

On its face, there is a reason why these FDA-regulated prescription drugs have a law enforcement agency heavily involved: legal controlled substances can be diverted, moved out of the legal supply chain into illicit use, whether stolen, resold, or obtained by forged prescription.

But placing regulatory oversight of licit controlled substance manufacturing in a law enforcement agency was a policy choice, not a necessity. The CSA, enacted in 1970, grew out of a drug-control framework advanced by the Nixon administration that placed law enforcement at the center of what it cast as a war on drugs. Three years later, Nixon folded its enforcement apparatus, along with several other Department of Justice (DOJ) drug units, into a single new agency built to run that war: the DEA.

DEA is built for enforcement: its mission is to investigate and prosecute those who traffic controlled substances, pursued with roughly 10,000 personnel. Regulating the legal supply of these drugs falls to one small office within the agency, the Diversion Control Division, whose staff are a small fraction of that total. That division's mandate, on paper, is dual: not only prevent diversion but also ensure an adequate and uninterrupted supply for legitimate medical, scientific, and commercial needs. The allocation and timing failures discussed above show how little the broader agency honors the second half of that division's mission. That imbalance is reflected in DEA's own mission statement, which emphasizes enforcement and prosecution while making no explicit mention of ensuring adequate supplies for medical use.

The policy response to the opioid epidemic reinforced the notion that access is not what DEA should pay attention to. As the epidemic wore on, Congress pressed DEA to cut quotas through hearings, letters, and eventually statute. Members faulted DEA for having let opioid production quotas climb during the epidemic and urged it to rein them in. The SUPPORT Act of 2018 turned that pressure into law, directing DEA to subtract an estimate of diversion from the opioid quota for five named opioids. But DEA took all these signals well beyond the text, treating supply as something to cut across the whole opioid category, and extending that logic to ADHD stimulants, which Congress had never named at all. (In fact, Congress originally scheduled ADHD drugs as Schedule III, not subject to quota, with DEA "up-scheduling" them over time.)

Diversion and inappropriate use have since often been invoked to justify tightening quota, but quota is the wrong tool for both. It acts at the manufacturing stage, while the conduct DEA seeks to change happens downstream, in prescribing, dispensing, and distribution. A ceiling cannot single out diversion or inappropriate use; it reaches only the total amount made. The relevant policy question is not how much medicine to withhold from the whole market, but how to identify and address the specific prescribing, dispensing, or diversion behaviors of concern.

For oral opioids, the meaningful reductions in inappropriate use came through demand-side measures: prescription drug monitoring programs, prescribing guidelines, and enforcement against pill mills. Those tools work by distinguishing appropriate prescribing from inappropriate prescribing, exactly the distinction a quota cannot make.

Yet quota remains attractive to DEA because it requires no such judgment. It offers a visible response to political pressure, measured in a published number and announced as a reduction in supply. An enforcement agency guards against one kind of error, letting a diverter through, and discounts the other, denying a patient. Faced with that trade-off, it will predictably keep reaching for the same cut, whatever the drug.

DEA's handling of a 2023 timing change shows the reflex with injectable opioids: it moved Schedule II quota to quarterly allocation over manufacturers' warnings that it would make production planning "extremely difficult if not untenable," implemented it anyway, and reversed course only as morphine and other injectable shortages loomed.

The problem is not one bad decision but an agency whose mission makes such decisions predictable.

Why FDA is the right regulator

There was a moment when licit drug supply regulation might have gone to health rather than law enforcement. In 1963, President Kennedy's Advisory Commission on Narcotic and Drug Abuse recommended a deliberate split: drug-trafficking investigation would go to the DOJ, while regulation of the licit manufacture and supply of these drugs, including the authority to set manufacturing quotas, would go to what became today's Department of Health and Human Services.

The Commission's reasoning was straightforward: deciding how much medicine the country needs and keeping it in adequate supply turns on clinical questions—how many patients, which conditions, what formulations—while pursuing those who divert it into illicit channels is police work. The two call for different expertise and serve different ends, so they belong in different hands.

The CSA reversed half of that design. Enacted in 1970, it honored the split on enforcement recommendation, sending trafficking investigation to what became the DEA, but it sent quota to the attorney general as well, placing regulation of the licit drug supply inside a law enforcement agency.

The Kennedy Commission's logic is as sound now as it was in 1963, and it points to FDA. FDA already sets the manufacturing standards these facilities must meet, inspects them for compliance, and approves the products they make. Quota is the one lever governing these same facilities that sits with a different agency, under a different mission. Moving it would close a gap in a job FDA already does rather than hand it an unfamiliar one.

FDA works to prevent shortages and to mitigate the ones it cannot prevent, but for drugs under DEA quota its toolkit runs into a wall. When a disruption hits, FDA's Drug Shortage Staff contacts the affected maker and its competitors, checks inventory, and asks who can raise output. For Schedule II drugs, the moves that would close the gap are the ones FDA cannot make: a competitor can raise output only with more quota and can hold the stock to stage that production only within DEA's inventory limits.

With control over quota, FDA could also make supply more reliable before any disruption occurs. It holds the plant-level quality and inspection record for every one of these facilities and requires makers of critical drugs to map their supply chain vulnerabilities, so it knows which are reliable, which are single-source, and which are fragile. A quota-setter with that record could steer production toward the makers most likely to deliver and away from concentrated, fragile sites, lowering the odds a disruption occurs at all.

Ensuring an adequate supply of generic medicines requires both sufficient quota and regulatory predictability. A manufacturer deciding whether to keep making a low-margin generic weighs not just the price but whether regulatory decisions are transparent and predictable: whether it will know how decisions are made, get reasons when an application is denied, and have a path to fix problems. FDA supplies that through published guidance and reasoned decisions, because managing an ongoing relationship with manufacturers is its core function; DEA's administration of quota offers little of it.

Steering production toward some manufacturers and away from others, as both reliability and access would require, raises two objections. First, FDA has long said it does not tell manufacturers what to make. But a quota is permission, not a command: a manufacturer applies for the quantity it wants to produce, and the regulator grants or denies it, exactly as FDA already does with every application to make a drug. Second, quota necessarily picks among firms. But FDA already does that for the sake of access, through published policy rather than ad hoc choice. In a shortage, it expedites the makers positioned to close the gap and extends them the flexibility it withholds from others. And under its first-generic policy, it prioritizes review of the earliest eligible applications where competition is thin.

Control over quota would do more than route supply to reliable makers; it would also let FDA steer it toward the versions patients most need. FDA cannot today influence which version a manufacturer makes; a firm holding one authorization channels it to the higher-margin brand or formulation, and the version patients depend on goes short. Allocating quota by version would change that: FDA could grant quota for the generic, or for the injectable in short supply rather than the patch, so the authorized supply matches the need. The firm still chooses whether to produce under that quota; FDA sets only what it is for.

Moving quota to FDA would do more than improve drug supply; it would make FDA accountable for balancing appropriate access against inappropriate use. If DEA holds the quota, FDA can treat inappropriate use as someone else's problem. Move the quota, and that changes: the pressure to curb inappropriate use would land on FDA. And because a quota cut denies treatment to appropriate patients along with the rest, FDA could not answer that pressure by cutting supply. It would have to use the tools it holds, particularly risk evaluation and mitigation strategies (REMS), which can impose conditions on prescribing and dispensing.

What stays with DEA

Moving responsibility for administering quotas to FDA would likely be good for DEA. Administering supply is a job that runs against the grain of an enforcement agency—it means housing a division whose goal, keeping drugs available, points the opposite way from the agency's mission of restricting their misuse. The transfer would let DEA focus on the enforcement work it is built for and stop it from answering for supply outcomes it was never equipped to deliver.

Even with quota administration moved to FDA, DEA would keep registration—the gatekeeping function that decides who may make, distribute, dispense, or prescribe a controlled substance. A manufacturer would hold its DEA registration while applying to FDA for quota, so FDA could grant quota only to a firm whose registration is current, and that quota would lapse automatically if DEA suspended or revoked registration.

DEA would also retain all its manufacturing-stage enforcement authorities. Regulation of facility-level diversion controls would stay with DEA: security and personnel requirements, recordkeeping, suspicious-order reporting, and the inspections behind them. The enforcement tools stay too: an Immediate Suspension Order to halt operations before any hearing when continued operation poses an imminent danger, revocation for cause through an Order to Show Cause, civil penalties, and binding

compliance terms through a Memorandum of Agreement. Every one of these controls and powers attaches to the manufacturer's DEA registration, which is why registration should also remain with DEA.

Nor would the transfer weaken diversion control. Diversion does not originate where the quota sits. It happens downstream, in distribution, at pharmacies, and in prescribing, not at the manufacturing stage a quota governs. DEA's authority over those downstream stages would not change: it would retain enforcement control over the distributors, pharmacies, and prescribers it registers today.

Notably, DOJ's Office of Inspector General did not identify manufacturing quota as a diversion-control tool. Although the sharp rise in opioid quota featured prominently in its review of DEA's handling of opioid diversion, none of its nine recommendations proposed using quota to control diversion. Instead, it pointed to registration, suspicious-order data, electronic prescribing, and interagency coordination, all aimed at the downstream stages where diversion occurs.

What moving quota to FDA requires

The statutory change that matters most is moving quota-setting from DEA to FDA—the health agency that approves these drugs, inspects the plants that make them, and answers for drug shortages. The other changes in this section would help whoever administers quota, but none substitutes for moving it to the agency whose mission fits the job.

Statutory changes to support FDA

Cost recovery. Congress should extend to FDA the cost-recovery authority that currently funds this work at DEA. Since 1993, the CSA has required DEA to set registration fees sufficient to recover the full cost of administering its diversion-control program, and the same authority should follow the function to FDA, sized to what it actually costs to administer quota well. The transfer burdens FDA with a function Congress has not separately funded only if Congress declines to move the fee with it.

DEA would give up little in the transfer. Manufacturer registration fees run $3,699 a year across roughly 560 registrants—about $2.1 million, a small fraction of the roughly $650 million Diversion Control Fee Account that funds the program. The function DEA would shed is a minor part of its resourcing, not a stake worth defending—and the modest fees suggest it has never been resourced as more than an afterthought.

Stranded quota. Congress should clarify that the agency in charge could reassign stranded quota. This clarification is needed because the statute spells out only two situations for readjusting quota that has already been granted: when a registrant loses its DEA registration, or when the aggregate is lowered for everyone. It is silent on the far more common case of a compliant manufacturer that cannot use its quota after a contamination event or equipment breakdown, and DEA has read that silence to mean it must wait for the firm to relinquish the quota. The fix is narrow: authority to reclaim demonstrably unusable quota and reallocate it, without cutting the surrendering firm's future baseline, so that giving up quota it cannot use costs the firm nothing.

Imports. Because DEA would keep control of imports after the transfer, two statutory changes are needed. First, Congress should set an import-authorization deadline. An importer cannot begin arranging the export permits that source countries such as Turkey and Australia require until it holds a U.S. authorization, and that process takes months, so a late authorization means the drug arrives too late for the year it was meant to supply.

Second, a statutory backstop should prevent DEA from using its import permit authority to undercut FDA's quota. FDA would set the total the country needs, but DEA would still issue the per-shipment

import permits. Without an obligation to support FDA's estimate, DEA could deny or delay those permits and hold actual imports below the level FDA found necessary, undoing the quota.

Downstream data. FDA would need better downstream data for both its current role in estimating medical need and the expanded role the transfer would give it in addressing inappropriate use. The data come in two kinds: DEA systems FDA should be given access to and external prescribing and utilization data.

DEA's own systems are the first kind. The Automation of Reports and Consolidated Orders System (ARCOS) already tracks distribution to the dispensing level, and the Controlled Substance Ordering System (CSOS) captures controlled substance orders transaction by transaction, in close to real time; both sit inside the registration infrastructure DEA keeps, so the ask is a statutory channel routing that information to FDA. Aside from informing demand estimates, these data could help FDA monitor shortages and inform risk management programs.

External data are the second. Using cost-recovery funding, FDA will be able to buy commercial datasets. But no commercial dataset is complete enough to support a national quota. Available prescribing data are samples, limited to the systems that feed them, while comprehensive e-prescribing data are not for sale. So this data deficit needs a CSA reporting mandate on the intermediaries that route e-prescriptions, de-identified and walled off from law enforcement, giving FDA the comprehensive read on prescribing that no purchase can provide.

Diversion subtraction. Congress should change what DEA's opioid diversion estimate does. The SUPPORT Act of 2018 directs DEA to estimate diversion and subtract it from the quota for five opioids, including two of the three injectable opioids now in shortage. The instruction sounds sensible, but medical need and diversion are separate quantities: an estimate of medical need does not include the diverted amount, so subtracting diversion comes out of patient supply, not diverted supply, setting the ceiling below what patients need.

The estimate itself is worth keeping, because knowing where and how much diversion occurs is useful. But instead of subtracting it from the manufacturing ceiling, the statute should require DEA to assess where diversion occurs and propose measures to address it there. That is the approach the DOJ inspector general urged: use the data DEA already holds to target diversion where it occurs, rather than cutting the supply reaching patients. That would hold DEA accountable for acting on diversion where it actually happens, without denying supply to the patients who depend on the drug.

Quota deadline. The manufacturer-quota deadline also sends the wrong signal, like the diversion subtraction just discussed. The 2018 SUPPORT Act moved the deadline for fixing each manufacturer's quota to December 1, presumably to give DEA more room to deliberate. But a quota is authorization to run a months-long process, so a number that arrives weeks before the production year cannot yield product at the start of it. Congress should move the deadline back to October 1, signaling that quota exists to give manufacturers enough lead time to have product ready when the year begins.

Statutory direction for FDA's discretion

Much of the current quota system reflects discretionary choices DEA has made over time rather than statutory requirements. Inventory caps, reserve expectations, the procurement and export split, and shortage-response practices should therefore not automatically carry over unchanged when quota moves to FDA.

Where the inherited regulations do need revisiting, the right design varies by drug, so Congress should not legislate fixed values. The inventory a hospital injectable can safely hold is not the inventory that makes sense for an oral opioid or a stimulant.

Congress should instead direct FDA to review the transferred guidance and regulations through a benefit-risk lens that weighs patient access against diversion risk, with authority to let the answer differ by drug. Reviewing a rulebook this large is the kind of open-ended work an agency leaves undone without a push, so the direction should carry a deadline and a report to Congress on how each rule was adjusted and why.

Conclusion

Resilience is the goal everywhere else in American drug policy, but for these medicines, DEA treats any slack as diversion waiting to happen. The result is a system built around friction: quota is split into many grants, released late, capped by inventory limits, and locked to the firm that holds it, making supply less responsive when disruptions occur.

DEA's zero-slack posture has consequences. In a recent survey of hospital pharmacy executives, 79% reported that shortages of injectable hydromorphone, morphine, and fentanyl had forced meaningful changes to patient care; more than half of clinicians reported avoidable adverse events they attributed to the shortages; and one in five clinicians reported medication errors or near misses, mainly from unfamiliar substitute drugs and difficult dose conversions. The gaps landed where they could least be absorbed: pediatric oncology, obstetric anesthesia, the operating room, and end-of-life care.

The stimulant shortage has also inflicted a toll. A Centers for Disease Control and Prevention (CDC) survey found that among the roughly one third of adults with ADHD who took stimulants in 2023, more than seven in ten had difficulty filling their prescriptions. Interrupted treatment is a clinical harm, not a logistical one: unmanaged ADHD impairs executive function. And the CDC warned that patients cut off from regulated supply may turn to an illegal market where counterfeit pills made to look like Adderall increasingly contain fentanyl.

These failures stem from a mismatch between mission and responsibility. FDA answers for drug shortages and patient access; DEA answers for diversion and criminal enforcement. Correcting that mismatch follows the ordinary logic of government: regulation and enforcement are separate functions. DOJ prosecutes those who defraud Medicare, but it does not run Medicare. The same principle should apply to the manufacture of controlled substances.

Moving quota to FDA will not end shortages. No regulator can prevent every contamination event or equipment failure. But it would make the system more resilient. FDA could weigh which manufacturers have stronger risk management—information it already requires—and place quota where production is most likely to hold up rather than where disruption is likeliest. When a shortage did hit, it could clear the government-built barriers that now block recovery: the capped supply, the stranded quota, the firms unable to step in.

Congress should make the transfer and the statutory changes that support it—funding, reallocation authority, the import and data provisions, an earlier quota deadline, and repeal of the diversion subtraction. The statutory fixes matter, but the transfer is what changes the mission behind every quota decision, and that is the change that counts.

DEA should keep chasing diversion. Deciding how much medicine may be made should belong to FDA.